JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 75, 149–163 (1980)

Approximation by Double Least-Squares Inverses

CHARLES K. CHUI*

Department of Mathematics, Texas A & M University, College Station, Texas 77843

Submitted by Ky Fan

1. INTRODUCTION

The method of least-squares inverse was first studied by E. A. Robinson in [8] to obtain a minimum-delay finite-length wavelet whose convolution with a given finite-length wavelet produces a best approximation to the unit spike (1, 0,..., 0) in the l^2 sequence norm. A detailed study of this subject can be found in Robinson's book [7, pp. 167–174]. In terms of polynomial approximation, this result can be stated as follows: *Let $P_n(z) = p_0 + \cdots + p_n z^n$ be a polynomial with real coefficients such that $p_0 \neq 0$. Then the polynomial $A_k(z) = a_0 + \cdots + a_k z^k$ with real coefficients so chosen that the L^2-norm on the unit circle*

$$\| 1 - A_k P_n \|_2 = \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} | 1 - A_k(e^{i\theta}) \, P_n(e^{i\theta})|^2 \, d\theta \right\}^{1/2} \tag{1.1}$$

is minimum for all possible choices of kth *degree polynomials with real coefficients, has the property that A_k is zero-free in the closed unit disc $| z | \leqslant 1$.* This result also holds for polynomials with complex coefficients, and a simple proof of this fact, using orthogonal polynomials, will be included in the next section. The relationship between A_k and orthogonal polynomials was observed in [3] where only real coefficients were considered. When complex coefficients are used, A_k will be called the *least-squares inverse of P_n in π_k*, the vector space of all polynomials of degree at most k over the field of complex numbers. Let B_n be the least-squares inverse of A_k in π_n. Then B_n will be called the *double least-squares inverse of P_n through π_k*. Hence, this process transforms P_n to another polynomial B_n in the same space π_n, such that $B_n(z) \neq 0$ for all z in the closed unit disc $| z | \leqslant 1$.

In the theory of digital signal processing and geophysical studies, the most important model of the transfer function is a rational function $H(z) =$

* Supported in part by the U.S. Army Research Office under Grant No. DAAG 29-78-0097.

0022-247X/80/050149-15$02.00/0

$Q_m(z)/P_n(z)$, where Q_m and P_n are (relatively prime) polynomials in π_m and π_n, respectively, such that $P_n(0) \neq 0$. If P_n is not a constant, then $H(z)$ is the transfer function of a recursive digital filter, or an autoregressive moving-average (ARMA) system. Many approximation methods are available to obtain $H(z)$ such that the corresponding magnitude spectrum $|H(e^{-i\omega})|$ has the desired properties of an ideal filter (cf. [2, 5, 6]). However, a realistic system must be stable; that is, the polynomial P_n must be zero-free in $|z| \leqslant 1$ (cf. [6, 9]). Therefore, the above transformation of P_n to B_n, the double least-squares inverse of P_n through π_k, seems to be an excellent method to transform a not necessarily stable system $Q_m(z)/P_n(z)$ to the stable system $Q_m(z)/B_n(z)$. In fact, Shanks [9, 10] observed that in the all-pole (or AR) system, that is, $H(z) = 1/P_n(z)$, the magnitude spectrum of the transformed system is a very good approximation of that of the original one. This observation, however, needs some modification as will be seen in Corollary 3.1 of Section 3. In practice, the usual choice of k is $k = n$ (cf. [6, 9, 10]).

It is very natural to ask if the above process preserves polynomials which are zero-free in the closed unit disc: That is, if $P_n(z) \neq 0$ in $|z| \leqslant 1$, is it necessarily true that $B_n = P_n$, where B_n is the double least-squares inverse of P_n through π_n? Even for $n = 1$, the answer is negative. Indeed, if $P_1(z) = z + 2$, then its double least-squares inverse B_1 through π_1 is $B_1(z) = (525/741)\, z + 3045/1482$. A class of examples for $n = 1$ will be given in Section 3.

A main result of this paper is that the transformation by double least-squares inverse "eventually" preserves polynomials which are zero-free in $|z| < 1$. More precisely, we will prove that *if $B_{n,k}$ is the double least-squares inverse of P_n through π_k, then $B_{n,k}$ converges to P_n, as k tends to infinity, if and only if P_n is zero-free in $|z| < 1$.* Since π_n is a finite-dimensional vector space, coefficient-wise convergence is equivalent to convergence in any L^p-norm. Hence, we will not specify the type of convergence here and in the next two sections. The proof of the above result will be included in Section 2. In view of this result, it is advisable to use very large values of k. It is clear that the method of finding least-squares inverses is linear, and in fact, the coefficient matrix is a positive definite Hermitian banded Toeplitz matrix with bandwidth equal to $2n + 1$ (cf. Section 2). It therefore takes very little computer time to invert matrices of this type even if the dimension is very large. The first half of Section 3 is devoted to the case $n = 1$ and a convergence result for $P_1(z) = \alpha - z$ with $|\alpha| < 1$ is obtained. This result leads to the formulation of the convergence theorem for any polynomial P_n with $P_n(0) \neq 0$. In Section 4, a generalized double least-squares inverse problem is discussed and examples are given, and in the final section some open problems will be posed.

2. Orthogonal Polynomials and Least-Squares Inverses

Let P_n be in π_n, the space of all polynomials of degree at most n, such that $P_n(0) \neq 0$. With this P_n, we define a measure

$$d\mu(\theta) = \frac{1}{2\pi} \mid P_n(e^{i\theta})|^2 \, d\theta, \tag{2.1}$$

and let $\phi_k \in \pi_k$, $k = 0, 1,...,$ be the orthonormal polynomials on the unit circle with respect to this measure, such that the leading coefficients are positive. Here, we have used the inner product

$$\langle f, g\rangle_\mu := \int_{-\pi}^{\pi} f(e^{i\theta}) \overline{g(e^{i\theta})} \, d\mu(\theta). \tag{2.2}$$

so that $\langle \phi_k, \phi_j\rangle_\mu = \delta_{kj}$. As usual, let $\|f\|_\mu = \langle f, f\rangle_\mu^{1/2}$, and $L^2(d\mu)$ be the Hilbert space of all functions f on the unit circle $|z| = 1$ such that $\|f\|_\mu < \infty$. Hence, the function $1/P_n$ is in $L^2(d\mu)$ with $\|1/P_n\|_\mu = 1$. Let $A_k \in \pi_k$ be chosen such that

$$\left\| \frac{1}{P_n} - A_k \right\|_\mu = \inf \left\{ \left\| \frac{1}{P_n} - A \right\|_\mu : A \in \pi_k \right\}. \tag{2.3}$$

Clearly, we have

$$\left\| \frac{1}{P_n} - A_k \right\|_\mu = \| 1 - A_k P_n \|_2,$$

where $\|\cdot\|_2$ is the L^2-norm on the unit circle with respect to the Lebesgue measure as defined in (1.1). Hence, A_k is the (unique) least-squares inverse of P_n in π_k. From (2.3), we have

$$\langle 1/P_n - A_k, z^j\rangle_\mu = 0 \tag{2.4}$$

for $j = 0,..., k$. Since

$$\left\langle \frac{1}{P_n}, z^j \right\rangle_\mu = \frac{1}{2\pi} \int_{-\pi}^{\pi} \overline{P_n(e^{i\theta})} \, e^{-ij\theta} \, d\theta$$

$$= \frac{1}{2\pi} \overline{\int_{-\pi}^{\pi} P_n(e^{i\theta}) \, e^{ij\theta} \, d\theta} = \overline{P_n(0)} \, \delta_{j0},$$

the linear system (2.4) becomes

$$\langle A_k, z^j\rangle_\mu = \overline{P_n(0)} \, \delta_{j0} \tag{2.5}$$

for $j = 0,\dots, k$. Let $A_k^*(z) = z^k \bar{A}_k(z^{-1})$ be the reciprocal polynomial with respect to A_k. Then (2.5) gives the linear equations

$$\langle A_k^*, z^j \rangle_\mu = 0 \tag{2.6}$$

for $j = 0,\dots, k-1$ and $\langle A_k^*, z^k\rangle_\mu = P_n(0)$. The orthogonality condition (2.6) implies that A_k^* is a constant multiple of ϕ_k, and since $P_n(0) \neq 0$, this constant is not equal to zero. Since ϕ_k has all its zeros in the open unit disc $|z| < 1$ (cf. Szegö [11, p. 292]), so does A_k^*. That is, A_k is zero-free in $|z| \leqslant 1$. This proves the following.

PROPOSITION 2.1. *Let $P_n \in \pi_n$ with $p_n(0) \neq 0$. Then its least-squares inverse A_k in π_k is a nonzero constant multiple of ϕ_k^*, the reciprocal polynomial with respect to the orthonormal polynomial ϕ_k. In particular, $A_k(z) \neq 0$ for $|z| \leqslant 1$.*

The linear system (2.5) gives a simple method for obtaining least-squares inverses. Indeed, if $A_k(z) = a_0 + \cdots + a_k z^k$, then (2.5) becomes $C\mathbf{a} = \mathbf{p}$ where $\mathbf{a} = [a_0, \dots, a_k]^{\mathrm{T}}$, $\mathbf{p} = [\overline{P(0)}, 0,\dots, 0]^{\mathrm{T}}$ and $C = [c_{jl}]$ is a positive definite Hermitian Toeplitz matrix which is also banded with bandwidth equal to $2n + 1$, given by

$$c_{jl} = \frac{1}{2\pi} \int_{-\pi}^{\pi} | P(e^{i\theta})|^2 \, e^{-i(j-l)\theta} \, d\theta, \tag{2.7}$$

$j, l = 0,\dots, k$. Here, and throughout, the superscript T denotes the transpose of a matrix. This system $C\mathbf{a} = \mathbf{p}$ of "normal equations" was also derived in [7, pp. 169–170] for the real case.

Now, let $B_n = B_{n,k}$ be the least-squares inverse of A_k in π_n. That is, $B_{n,k}$ is the double least-squares inverse of P_n through π_k. Hence, $B_{n,k}(z) \neq 0$ for $|z| \leqslant 1$. If $B_{n,k}$ converges to P_n as $k \to \infty$, then $P_n(z) \neq 0$ for $|z| < 1$ since $P_n(0) \neq 0$ and the convergence is uniform. The main result in this section is to establish the converse. That is, we have the following.

THEOREM 2.1. *Let $P_n \in \pi_n$ with $P_n(0) \neq 0$, and $B_n = B_{n,k}$ be the double least-squares inverse of P_n through π_k. Then $B_{n,k} \to P_n$ as $k \to \infty$ if and only if $P_n(z) \neq 0$ for $|z| < 1$.*

If P_n is zero-free in the open unit disc $|z| < 1$, then $\log |P_n|$ is a harmonic function there. Hence, the geometric mean of $|P_n|$ on the unit circle is

$$\exp \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \log |P_n(e^{i\theta})| \, d\theta \right\} = |P(0)|,$$

even if P_n may have zeros on $|z| = 1$.

By a well-known result on orthogonal polynomials [11, p. 303], we have

$$\sum_{j=0}^{\infty} |\phi_j(0)|^2 = \frac{1}{|P_n(0)|^2}. \tag{2.8}$$

Since $1/P_n$ is analytic in $|z| < 1$ and is in $L^2(d\mu)$, it is the limit in $L^2(d\mu)$ of its Fourier series

$$\frac{1}{P_n} = \sum_{j=0}^{\infty} \left\langle \frac{1}{P_n}, \phi_j \right\rangle_\mu \phi_j .$$

Also, by (2.3), A_k is the best approximant of $1/P_n$ from π_k in $L^2(d\mu)$, so that

$$A_k = \sum_{j=0}^{k} \left\langle \frac{1}{P_n}, \phi_j \right\rangle_\mu \phi_j .$$

Hence, we have

$$\begin{aligned} \| 1 - A_k P_n \|_2^2 &= \left\| \frac{1}{P_n} - A_k \right\|_\mu^2 \\ &= \sum_{j=k+1}^{\infty} \left| \left\langle \frac{1}{P_n}, \phi_j \right\rangle_\mu \right|^2 \\ &= \sum_{j=k+1}^{\infty} \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} \overline{P_n(e^{i\theta})}\, \overline{\phi_j(e^{i\theta})}\, d\theta \right|^2 \\ &= \sum_{j=k+1}^{\infty} \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} (P_n \phi_j)(e^{i\theta})\, d\theta \right|^2 \\ &= |P_n(0)|^2 \sum_{j=k+1}^{\infty} |\phi_j(0)|^2 \to 0, \end{aligned} \tag{2.9}$$

as $k \to \infty$ by (2.8). Next, since $B_{n,k}$ is the least-squares inverse of A_k in π_n and $P_n \in \pi_n$, we have

$$\begin{aligned} \| P_n - B_{n,k} \|_2 &\leqslant \|(1 - B_{n,k} A_k) P_n \|_2 + \|(1 - A_k P_n) B_{n,k} \|_2 \\ &\leqslant \| 1 - B_{n,k} A_k \|_2 \| P_n \|_\infty + \| 1 - A_k P_n \|_2 \| B_{n,k} \|_\infty \\ &\leqslant \| 1 - A_k P_n \|_2 (\| P_n \|_\infty + \| B_{n,k} \|_\infty), \end{aligned} \tag{2.10}$$

where $\| \cdot \|_\infty$ is the supremum norm on the unit circle. Because all L^p-norms in π_n are equivalent, there is a positive constant c such that $\| D \|_\infty \leqslant c \| D \|_2$ for all $D \in \pi_n$. Hence, by using (2.10), we have

$$\begin{aligned} \| B_{n,k} \|_\infty &\leqslant c \| B_{n,k} \|_2 \leqslant c \| P_n \|_2 + c \| P_n - B_{n,k} \|_2 \\ &\leqslant c \| P_n \|_2 + c \| 1 - A_k P_n \|_2 (\| P_n \|_\infty + \| B_{n,k} \|_\infty), \end{aligned}$$

which gives

$$(1 - c\,\| 1 - A_k P_n \|_2)\,\| B_{n,k} \|_\infty \leqslant c\,\| P_n \|_2 + c\,\| 1 - A_k P_n \|_2\,\| P_n \|_\infty .$$

By applying (2.9), we obtain

$$\limsup_{k \to \infty} \| B_{n,k} \|_\infty \leqslant c\,\| P_n \|_2 .$$

Combining this with (2.10), we can find a positive constant d, such that

$$\| P_n - B_{n,k} \|_2 \leqslant d\,\| 1 - A_k P_n \|_2 . \tag{2.11}$$

In view of the estimate in (2.9), we have proved that

$$B_{n,k} \to P_n$$

as $k \to \infty$.

Note that (2.9) and (2.11) together give the order of approximation of P_n by its double least-squares inverse $B_{n,k}$ through π_k.

3. Convergence of Double Least-Squares Inverses

We will first discuss the case when $n = 1$. The result obtained in this case will help us to find a natural candidate for the limit function of the double least-squares inverses $B_{n,k}$ of P_n for any n. Without loss of generality, we consider

$$P_1(z) = \alpha - z,$$

where α is a nonzero complex number. Let

$$A_k(z) = a_0 + a_1 z + \cdots + a_k z^k$$

be the least-squares inverse of P_1 in π_k, and write $C\mathbf{a} = \mathbf{p}$ with $\mathbf{a} = [a_0, \ldots, a_k]^{\mathrm{T}}$ and $\mathbf{p} = [\bar{\alpha}, 0]^{\mathrm{T}}$ as in the previous section. It is clear that the coefficient matrix C is given by the tridiagonal Toepliz matrix

$$C = \begin{bmatrix} 1 + |\alpha|^2 & -\alpha & 0 & \cdots & 0 & 0 & 0 \\ -\bar{\alpha} & 1 + |\alpha|^2 & -\alpha & \ddots & & & 0 \\ & \ddots & \ddots & \ddots & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \ddots & -\alpha \\ 0 & \cdots & \cdots & 0 & & -\bar{\alpha} & 1 + |\alpha|^2 \end{bmatrix}.$$

Hence, it follows easily that

$$A_k(z) = \frac{\bar{\alpha}}{1 + \cdots + |\alpha|^{2k+2}} \times [(1 + \cdots + |\alpha|^{2k}) + \bar{\alpha}(1 + \cdots + |\alpha|^{2k-2})\, z + \cdots + \bar{\alpha}^k z^k]. \tag{3.1}$$

The least-squares inverse

$$B_{1,k}(z) = b_{0,k} - b_{1,k} z$$

of A_k in π_1, or equivalently the double least-squares inverse of P_1, can be determined in the same straightforward manner. To simplify the following presentation and to enable us to prove a convergence result to be stated later, we introduce the notation

$$d = |\alpha|^2 \qquad \text{and} \qquad c_j = \sum_{l=0}^{j} d^l. \tag{3.2}$$

Then we have the linear system

$$\frac{d}{c_{k+1}^2} \begin{bmatrix} \sum_{j=0}^{k} d^{k-j} c_j^2 & \alpha \sum_{j=1}^{k} d^{k-j} c_{j-1} c_j \\ \bar{\alpha} \sum_{j=1}^{k} d^{k-j} c_{j-1} c_j & \sum_{j=0}^{k} d^{k-j} c_j^2 \end{bmatrix} \begin{bmatrix} b_{0,k} \\ b_{1,k} \end{bmatrix} = \begin{bmatrix} \dfrac{\alpha c_k}{c_{k+1}} \\ 0 \end{bmatrix},$$

from which it follows that

$$b_{0,k}(\alpha) \equiv b_{0,k} = \frac{\alpha c_k c_{k+1} \sum_{j=0}^{k} d^{k-j} c_j^2}{d\left[\left(\sum_{j=0}^{k} d^{k-j} c_j^2\right)^2 - d\left(\sum_{j=1}^{k} d^{k-j} c_{j-1} c_j\right)^2\right]}, \tag{3.3}$$

$$b_{1,k}(\alpha) \equiv b_{1,k} = -\frac{c_k c_{k+1} \sum_{j=1}^{k} d^{k-j} c_{j-1} c_j}{\left[\left(\sum_{j=0}^{k} d^{k-j} c_j^2\right)^2 - d\left(\sum_{j=1}^{k} d^{k-j} c_{j-1} c_j\right)^2\right]}. \tag{3.4}$$

The only simple case here is when $|\alpha| = 1$, giving $d = 1$, $c_j = j + 1$, and

$$B_{1,k}(z) = \frac{4k+6}{4k+3}\alpha - \frac{4k}{4k+3} z.$$

Hence, for $|\alpha| = 1$, the zero of $B_{1,k}$ is at $(1 + 3/2k)\,\alpha$ and $B_{1,k} \to P_1$ as asserted by Theorem 2.1. In general, even though the expressions in (3.3) and (3.4) are quite complicated and cannot be simplified, Theorem 2.1 guarantees that both

$$(1/\alpha)\, b_{0,k}(\alpha) \to 1 \qquad \text{and} \qquad b_{1,k}(\alpha) \to 1, \tag{3.5}$$

for $|\alpha| \geqslant 1$, as $k \to \infty$. This fact also allows us to study the convergence of $B_{1,k}$ for $|\alpha| < 1$. Indeed, if $d = |\alpha|^2 < 1$, then setting

$$f = d^{-1} \qquad \text{and} \qquad g_j = \sum_{l=0}^{j} f^l, \tag{3.6}$$

we have

$$d = f^{-1} \qquad \text{and} \qquad c_j = f^{-j} g_j\,. \tag{3.7}$$

By using (3.7), the right-hand side of (3.3) becomes

$$b_{0,k} = \frac{\alpha f^{-k} g_k f^{-k-1} g_{k+1} \sum_{j=0}^{k} f^{-k+j} f^{-2j} g_j^2}{f^{-1}\left[\left(\sum_{j=0}^{k} f^{-k+j} f^{-2j} g_j^2\right)^2 - f^{-1}\left(\sum_{j=1}^{k} f^{-k+j} f^{-2j+1} g_{j-1} g_j\right)^2\right]}$$

$$= \frac{\alpha g_k g_{k+1} \sum_{j=0}^{k} f^{k-j} g_j^2}{\left[\left(\sum_{j=0}^{k} f^{k-j} g_j^2\right)^2 - f\left(\sum_{j=1}^{k} f^{k-j} g_{j-1} g_j\right)^2\right]}, \tag{3.8}$$

and similarly, the right-hand side of (3.4) can be written as

$$b_{1,k} = -\frac{g_k g_{k+1} \sum_{j=1}^{k} f^{k-j} g_{j-1} g_j}{\left[\left(\sum_{j=0}^{k} f^{k-j} g_j^2\right)^2 - f\left(\sum_{j=1}^{k} f^{k-j} g_{j-1} g_j\right)^2\right]}. \tag{3.9}$$

Observe that the expressions in the right-hand sides of (3.8) and (3.9) are the same as those of (3.3) and (3.4) except that in (3.8) we need the factor $1/f$. Furthermore, the relationship between g_j and f in (3.6) is identical to that between c_j and d in (3.2). Hence, for $|\alpha| < 1$ or equivalently, $f > 1$, the convergence result in (3.5) gives

$$(1/\alpha)\, b_{0,k}(\alpha) \to f \qquad \text{and} \qquad b_{1,k}(\alpha) \to 1.$$

Combining this result with (3.5), we have

$$\begin{aligned} b_{0,k}(\alpha) &\to \alpha && \text{if} \quad |\alpha| \geqslant 1, \\ &\to 1/\bar{\alpha} && \text{if} \quad 0 < |\alpha| < 1; \\ b_{1,k}(\alpha) &\to 1 \end{aligned} \tag{3.10}$$

for all nonzero complex number α. That is, we have proved the following:

PROPOSITION 3.1. *Let* $P_1(z) = \alpha - z$ *with* $0 < |\alpha| \leqslant 1$, *and* $B_{1,k}$ *be its double least-squares inverse through* π_k. *Then*

$$B_{1,k}(z) \to (1/\bar{\alpha} - z), \tag{3.11}$$

as $k \to \infty$.

Of course, for $|\alpha| \geqslant 1$ Theorem 2.1 guarantees that the double least-squares inverse $B_{1,k}$ of $P_1(z) = \alpha - z$ "eventually" preserves P_1, but (3.11) says that for $0 < |\alpha| < 1$, the double least-squares inverse "eventually" moves the zero of P_1 to its image of reflection across the unit circle. This result allows us to choose the natural candidate for the limit function of the sequence of double least-squares inverse for any polynomial P_n with $P_n(0) \neq 0$. Namely, we establish the following result.

THEOREM 3.1. *Let* $P_n \in \pi_n$ *be written as*

$$P_n(z) = (\alpha_1 - z) \cdots (\alpha_m - z) \cdot Q_{n-m}(z)$$

where $Q_{n-m} \in \pi_{n-m}$ *is zero-free in* $|z| < 1$ *and* $0 < |\alpha_j| < 1$ *for* $j = 1,..., m$, *and denote by* $B_{n,k} \in \pi_n$ *the double least-squares inverse of* P_n *through* π_k. *Then* $B_{n,k} \to \tilde{P}_n$ *as* $k \to \infty$, *where*

$$\tilde{P}_n(z) = (1/\bar{\alpha}_1 - z) \cdots (1/\bar{\alpha}_m - z) \cdot Q_{n-m}(z).$$

We will present an elegant proof of this result by Dr. E. T. Y. Lee. But first let us mention a consequence of this theorem. We observe that the magnitude spectrum corresponding to $B_{n,k}$ is *not* the same as that corresponding to the original polynomial P_n, for very large values of k. In fact, we have the following:

COROLLARY 3.1. *Let* P_n *and* $B_{n,k}$ *be defined as in Theorem* 3.1. *Then*

$$\lim_{k\to\infty} |B_{n,k}(e^{i\omega})| = c\, |P_n(e^{i\omega})|,$$

for all real ω, *where* $c = |\alpha_1 \cdots \alpha_m|^{-1}$.

Note that if P_n has zeros in $|z| < 1$, then $c > 1$, so that an adjustment of a multiplicative constant is needed to preserve the magnitude spectrum. That is, Shanks' procedure [10] requires a modification. We now give a proof of Theorem 3.1. In view of Theorem 2.1, it is sufficient to show that P_n and $\tilde{P}_n$ defined in the statement of the theorem have the same least-squares inverse in π_k. But this is trivial. Indeed, if we consider the measures $d\mu = (1/2\pi)\, |P(e^{i\theta})|^2\, d\theta$ and $d\tilde{\mu} = (1/2\pi)\, |\tilde{P}(e^{i\theta})|^2\, d\theta$, then it is clear that

$$d\mu = \gamma\, d\tilde{\mu}, \tag{3.13}$$

where $\gamma = |\alpha_1 \cdots \alpha_m|^2$. At the same time, we also have

$$P_n(0) = \gamma \tilde{P}_n(0). \tag{3.14}$$

Hence, if $A_k(z) = a_0 + \cdots + a_k z^k$ and $\tilde{A}_k(z) = \tilde{a}_0 + \cdots + \tilde{a}_k z^k$ denote the least-squares inverses of P_n and $\tilde{P}_n$ in π_k, respectively, then we have

$$C[a_0, \ldots, a_k]^{\mathrm{T}} = [\overline{P_n(0)}, 0, \ldots, 0]^{\mathrm{T}},$$

$$\tilde{C}[\tilde{a}_0, \ldots, \tilde{a}_k]^{\mathrm{T}} = [\overline{\tilde{P}_n(0)}, 0, \ldots, 0]^{\mathrm{T}},$$

where, because of (2.7) and (3.13), the coefficient matrices C and $\tilde{C}$ satisfy the relationship $C = \gamma\tilde{C}$. Hence, using (3.14) and the nonsingularity of C, we conclude that $[a_0, \ldots, a_k] = [\tilde{a}_0, \ldots, \tilde{a}_k]$, or $A_k = \tilde{A}_k$.

4. Generalized Double Least-Squares Inverses

Let H^2 be the usual Hardy space of functions f analytic in the open unit disc $|z| < 1$ with

$$\|f\|_2 = \sup_{0 \leqslant r < 1} \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(re^{i\theta})|^2 \, d\theta \right\}^{1/2} < \infty.$$

If $f \in H^2$ with $f(0) \neq 0$, $B_{n,k} \in \pi_n$ will be called the ***generalized double least-squares inverse of f in π_n*** through π_k, provided $B_{n,k}$ is the least-squares inverse of $A_k \in \pi_k$ in π_n, where A_k is the least-squares inverse of f in π_k. By using the same argument given in Section 2, we note that $A_k(z) \neq 0$ for all $|z| \leqslant 1$. Also, since H^2 is a subclass of the Nevanlinna class N, we conclude as in Section 2 that if $f \in H^2$ is zero-free in $|z| < 1$, and $\phi_j \in \pi_j$ are the orthonormal polynomials with respect to the measure $(1/2\pi)|\tilde{f}(\theta)|^2 \, d\theta$ on $|z| = 1$, having positive leading coefficients, then

$$\sum_{j=0}^{\infty} |\phi_j(0)|^2 = \frac{1}{|f(0)|^2} \tag{4.1}$$

(cf. [11, p. 303]). Here, and throughout, $\tilde{f}(\theta)$ denotes the almost everywhere radial limit of $f(re^{i\theta})$ as $r \uparrow 1$. That $\tilde{f}$ exists is due to the well-known theorem of Fatou. If, in addition, A_k is the least-squares inverse of f in π_k, then the same argument given in Section 2 yields

$$\| 1 - A_k f \|_2 = |f(0)| \left\{ \sum_{j=k+1}^{\infty} |\phi_j(0)|^2 \right\}^{1/2}, \tag{4.2}$$

which converges to zero as $k \to \infty$ by (4.1). Now, let $B_{n,k} \in \pi_n$ be the generalized double least-squares inverse in π_n of f through π_k; that is, $B_{n,k}$ is the least-squares inverse of A_k in π_n. It is natural to study the convergence of $B_{n,k}$ as $k \to \infty$, and if $B_{n,k} \to B_n$, to study how good is B_n an approximant of f from π_n. Let us start with two examples.

If $f(z) = 1 - z$, then the generalized double least-squares inverse of f in π_0 through π_k is $B_{0,k}(z) = 6/(2k+3)$. Hence, $B_{0,k} \to B_0 \equiv 0$ as $k \to \infty$.

Next, let $f(z) = 1 - z^2$, and $A_k(z) = \alpha_0 + \alpha_1 z + \cdots + \alpha_k z^k$ be its least-squares inverse in π_k. Then $C[\alpha_0, \ldots, \alpha_k]^{\mathrm{T}} = [1, 0, \ldots, 0]^{\mathrm{T}}$, where

$$C = \begin{bmatrix} 2 & 0 & -1 & & & \bigcirc \\ 0 & \ddots & \ddots & \ddots & & \\ -1 & \ddots & \ddots & \ddots & \ddots & \\ & \ddots & \ddots & \ddots & \ddots & -1 \\ & & \ddots & \ddots & \ddots & 0 \\ \bigcirc & & & -1 & 0 & 2 \end{bmatrix}.$$

A simple manipulation gives

$$\begin{aligned} A_k(z) &= \frac{1}{k+1}\left[k + (k-1)\,z^2 + (k-2)\,z^4 + \cdots + \frac{k}{2}\,z^k\right] && \text{if } k \text{ is even,} \\ &= \frac{1}{k+1}\left[k + (k-1)\,z^2 + (k-2)\,z^4 + \cdots + \frac{k+1}{2}\,z^{k-1}\right] && \text{if } k \text{ is odd} \end{aligned}$$

and the generalized double least-squares inverse $B_{1,k}$ of f in π_1 through π_k is

$$\begin{aligned} B_{1,k}(z) &= k(k+1)/[k^2 + \cdots + (k/2)^2] && \text{if} \quad k \text{ is even,} \\ &= k(k+1)/[k^2 + \cdots + ((k+1)/2)^2] && \text{if} \quad k \text{ is odd.} \end{aligned}$$

Hence, we have $B_{1,k}(z) \to B_1(z) \equiv 0$ as in the first example.

The above examples indicate that the following conjecture should hold.

CONJECTURE. *If f is a polynomial with m zeros (counting multiplicities) on $|z| = 1$ and is zero-free in $|z| < 1$, and if $B_{n,k}$ is its generalized double least-squares inverse in π_n through π_k, then $B_{n,k} \to 0$ as $k \to \infty$ for $n = 0, \ldots, m-1$.*

If the polynomial f is zero-free in $|z| \leqslant 1$, one would expect $B_{n,k} \to B_n$, the least-squares inverse of $1/f$ in π_n. This, and more, is true as in the following.

THEOREM 4.1. *Let $f \in H^2$ be such that its reciprocal $1/f$ is in H^∞, and $B_{n,k}$ be its generalized double least-squares inverse in π_n through π_k. Then $B_{n,k} \to B_n$ as $k \to \infty$ where $B_n \in \pi_n$ is the least-squares inverse of $1/f$ in π_n.*

If we set $d\nu = (1/2\pi)\,|\tilde{f}(\theta)|^{-2}\,d\theta$, where $\tilde{f}(\theta)$ is the almost everywhere radial limit of $f(re^{i\theta})$ as $r \uparrow 1$, and define, as before, the inner product

$$\langle g, h\rangle_\nu = \int_{-\pi}^{\pi} \tilde{g}\bar{\tilde{h}}\,d\nu,$$

then B_n is the best approximant of f from π_n in $\|\cdot\|_\nu = \langle\cdot,\cdot\rangle_\nu^{1/2}$. Let $\psi_j \in \pi_j$ be the orthonormal polynomials with positive leading coefficients with respect to this inner product. Then as in (4.1), we have

$$\sum_{j=0}^{\infty} |\psi_j(0)|^2 = |f(0)|^2,$$

$$\begin{aligned}\|B_n - f\|_\nu &= \inf\{\|B - f\|_\nu : B \in \pi_n\} \\ &= \|1 - B_n/f\|_2 = |f(0)|^{-1}\left\{\sum_{j=n+1}^{\infty} |\psi_j(0)|^2\right\}^{1/2} \to 0\end{aligned} \tag{4.3}$$

as $n \to \infty$.

To prove the theorem, let A_k be the least-squares inverse of f in π_k. Then A_k is zero-free in $|z| \leqslant 1$. Set $d\nu_k = (1/2\pi)\,|A_k(e^{i\theta})|^2\,d\theta$ and let $\langle\ ,\ \rangle_{\nu_k}$ be the corresponding analogous inner product. Also, let $\psi_{j,k} \in \pi_j$ be the orthonormal polynomials with positive leading coefficients with respect to $\langle\ ,\ \rangle_{\nu_k}$. Then

$$\frac{1}{A_k(z)} = \sum_{j=0}^{\infty} \left\langle \frac{1}{A_k}, \psi_{j,k}\right\rangle_{\nu_k} \psi_{j,k}(z)$$

and hence

$$\begin{aligned}B_{n,k}(z) &= \sum_{j=0}^{n} \left\langle \frac{1}{A_k}, \psi_{j,k}\right\rangle_{\nu_k} \psi_{j,k}(z) \\ &= \overline{A_k(0)} \sum_{j=0}^{n} \overline{\psi_{j,k}(0)}\,\psi_{j,k}(z).\end{aligned}$$

But since A_k is the least-squares inverse of f in π_k, we have

$$\begin{aligned}|1 - A_k(0)f(0)| &\leqslant \frac{1}{2\pi}\int_{-\pi}^{\pi} |1 - A_k(e^{i\theta})\tilde{f}(\theta)|\,d\theta \\ &\leqslant \|1 - A_k f\|_2 \to 0\end{aligned} \tag{4.4}$$

as $k \to \infty$, by (4.2). This gives $\overline{A_k(0)} \to \overline{f(0)^{-1}}$. To study the convergence of the orthogonal polynomials $\psi_{j,k}$ as $k \to \infty$, we recall that

$$\psi_{j,k}(z) = (D_{j-1,k}D_{j,k})^{-1/2} \begin{vmatrix} c_{0,k} & c_{-1,k} & \cdots & c_{-j,k} \\ c_{1,k} & c_{0,k} & \cdots & c_{-j+1,k} \\ \cdots & \cdots & \cdots & \cdots \\ c_{j-1,k} & c_{j-2,k} & \cdots & c_{-1,k} \\ 1 & z & \cdots & z^j \end{vmatrix}$$

where

$$c_{l,k} = \frac{1}{2\pi} \int_{-\pi}^{\pi} |A_k(e^{i\theta})|^2 e^{-il\theta} \, d\theta,$$

$l = 0, \pm 1, \ldots$ and $D_{l,k}$ is the (positive) determinant

$$D_{l,k} = \begin{vmatrix} c_{0,k} & c_{-1,k} & \cdots & c_{-l,k} \\ c_{1,k} & c_{0,k} & \cdots & c_{-l+1,k} \\ \cdots & \cdots & \cdots & \cdots \\ c_{l,k} & c_{-1,k} & \cdots & c_{0,k} \end{vmatrix}$$

(cf. [11, p. 288]). The orthogonal polynomial ψ_j can be written in the same manner with $c_{l,k}$ replaced by c_l. Hence, to prove that $\psi_{j,k} \to \psi_j$ as $k \to \infty$, it is sufficient to prove that $c_{l,k} \to c_l$, where

$$c_l = \frac{1}{2\pi} \int_{-\pi}^{\pi} |\tilde{f}(\theta)|^{-2} e^{-il\theta} \, d\theta$$

for each $l = 0, \pm 1, \ldots$. That is, we must show that $d\nu_k$ converges to $d\nu$ in the weak topology. Since $A_k \to 1/f$ uniformly on compact subsets of $|z| < 1$, it is equivalent to show that $\| A_k - 1/f \|_2 \to 0$. This is clear since $1/f \in H^\infty$ and

$$\left\| A_k - \frac{1}{f} \right\|_2 = \left\| (1 - A_k f) \frac{1}{f} \right\|_2 \leqslant \left\| \frac{1}{f} \right\|_\infty \| 1 - A_k f \|_2,$$

which tends to zero by (4.2). Hence, together with (4.4), we have

$$\begin{aligned} B_{n,k}(z) &= \overline{A_k(0)} \sum_{j=0}^{n} \overline{\psi_{j,k}(0)} \, \psi_{j,k}(z) \to \overline{f(0)}^{-1} \sum_{j=1}^{n} \overline{\psi_j(0)} \, \psi_j(z) \\ &= B_n(z). \end{aligned}$$

This completes the proof of the theorem.

Let f be analytic in $|z| < 1$ such that its reciprocal $1/f$ is in H^2, and let $B_n \in \pi_n$ be the least-squares inverse of $1/f$ in π_n. Then B_n is zero-free in $|z| \leqslant 1$. As in Section 2, let B_n^* denote the reciprocal polynomial with respect to B_n. Then by a well-known result of I. Schur, we have

$$|B_n^*(z)| \leqslant |B_n(z)| \tag{4.5}$$

for all z, $|z| \leqslant 1$, and that for each $m = 0, 1, 2, \ldots$,

$$P_{n,m}(z) = B_n(z) + z^m B_n^*(z) \tag{4.6}$$

is a polynomial in π_{m+n} with all its zeros lying on the unit circle C: $|z| = 1$. Such polynomials are called C-polynomials. For more information, see [1] and the references therein. For $z = re^{i\theta}$, $0 < r < 1$, we have

$$\begin{aligned} |1 - B_n(z)/f(z)| &\leqslant \frac{1}{2\pi} \int_{-\pi}^{\pi} |1 - B_n(e^{it})/\tilde{f}(t)|\, P_r(\theta - t)\, dt \\ &\leqslant C_r \left\| 1 - B_n \frac{1}{f} \right\|_2, \end{aligned}$$

where $P_r(\theta)$ is the Poisson kernel. Since B_n is the least-squares inverse of $1/f$, we have $\| 1 - B_n/f \|_2 \to 0$ by (4.2). That is, $B_n \to f$ uniformly on compact subsets of $|z| < 1$. Hence, by (4.5), we have $z^m B_n^*(z) \to 0$ uniformly on each compact subset of $|z| < 1$ as $m \to \infty$ independent of $n = 0, 1, 2, \ldots$. Therefore, the C-polynomials $P_{n,m}$ converge to f uniformly on compact subsets of $|z| < 1$ as both m and $n \to \infty$ independently. Since B_n is obtained from f by a linear method, this gives a straightforward method for constructing C-polynomials that approximate f uniformly on compact subsets of $|z| < 1$. We remark that uniform approximation on compact subsets of $|z| < 1$ cannot be replaced by approximation in the H^2-norm, even if f is in H^2. This was proved in [1, Theorem 2].

5. Final Remarks

There are many interesting and perhaps very difficult problems one can pose related to the results and discussion in this paper. It is hoped that this article stimulates a new area of research in digital filtering and approximation theory. In the case of one complex variable as discussed here, the conjecture given in Section 4 should be settled, and the analogous problems in H^p, $p \neq 2$, seem very difficult. Of course, these problems can be posed in the case of several complex variables. For example, Shanks [10] conjectured that the least-squares inverses of polynomials of two complex variables are zero-free in the bi-disc $\{(z_1, z_2): |z_1| < 1, |z_2| < 1\}$. This conjecture was, however, shown to be false by Genin and Kamp in [3]. Under certain restrictions, Shanks' conjecture should hold and this would be a very important tool in two-dimensional recursive digital filter design. A result analogous to Theorem 2.1 is important to "preserve" a stable filter under double least-squares inverses. For the approximation theorist, analogous problems can be posed on an interval, say [0, 1], of the real line. If $\| \cdot \|_p$ denotes the L^p-norm on [0, 1], $1 \leqslant p \leqslant \infty$, one can study (best) L^p-inverses and (best) double L^p-inverses analogous to the least-squares and double least-squares inverses, respectively, as discussed in this article.

ACKNOWLEDGMENT

We wish to thank Dr. E. T. Y. Lee for his elegant proof of Theorem 3.1.

REFERENCES

1. C. K. CHUI, C-polynomial approximation of H^p and $\mathscr{H}^p$ functions, *J. Math. Anal. Appl.* **34** (1971), 82–85.
2. C. K. CHUI AND A. K. CHAN, A two-sided rational approximation method for recursive digital filtering, *IEEE Trans. ASSP* **27** (1979), 141–145.
3. Y. GENIN AND Y. KAMP, Counter-examples in the best least-squares inverse stabilization of 2D-recursive filters, *Elec. Lett.* **11** (1975), 330–331.
4. THOMAS S. HUANG, Stability of two-dimensional recursive filters, *in* "Selected Papers in Digital Processing," Vol. II, pp. 290–295, Digital Signal Processing Committee, IEEE ASSP Soc., IEEE, New York, 1975.
5. D. L. MOFFATT AND K. A. SHUBERT, Natural resonances via rational approximants, *IEEE Trans. Antennas and Propagation* **25** (1977), 657–660.
6. A. V. OPPENHEIM AND R. W. SCHAFER, Digital Signal Processing, Prentice–Hall, Englewood Cliffs, N. J., 1975.
7. E. A. ROBINSON, "Statistical Communication and Detection," Hafner, New York, 1967.
8. E. A. ROBINSON, Structural properties of stationary stochastic processes with applications, *in* "Time Series Analysis" (M. Rosenblatt, Ed.), Wiley, New York, 1963.
9. J. L. SHANKS, Recursion for digital processing, *Geophysics* **32** (1967), 33–51.
10. J. L. SHANKS, The design of stable two-dimensional recursive filters, *in* "Proceedings, UMR – M. J. Kelly Commun. Conf.," Univ of Missouri, Rolla, 1970.
11. G. SZEGÖ, "Orthogonal Polynomials," *Amer. Math. Soc. Colloq. Publ.*, Vol. **23**, Providence, R. I., 1939.